



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

REC'D 28 JUN 2000

PC

POT

Bescheinigung

Certificate

Attestation

JKU

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

99306931.9

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE, 23/05/00
LA HAYE, LE

THIS PAGE BLANK (USPTO)



Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 99306931.9
Demande n°:

Anmeldetag:
Date of filing: 31/08/99
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
LUCENT TECHNOLOGIES INC.
Murray Hill, New Jersey 07974-0636
UNITED STATES OF AMERICA

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:

Method and apparatus for macroblock DC and AC coefficient prediction for video coding

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:
State:
Pays:

Tag:
Date:
Date:

Aktenzeichen:
File no.
Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:

H04N7/26, H04N7/30

Am Anmeldetag benannte Vertragsstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du dépôt:

Bemerkungen:
Remarks:
Remarques:

The original title of the application reads as follows:
Apparatus for compressing and expanding video data

THIS PAGE BLANK (USPTO)

APPARATUS FOR COMPRESSING AND EXPANDING VIDEO DATA

This invention relates to apparatus for compressing and expanding video data.

Existing video compression standards are all based on block discrete cosine
5 transform (DCT) transform. The picture is divided into square blocks consisting of 8x8
pixels. The blocks may contain the actual pixels or the prediction residual, which is the
difference between the actual and motion compensated block pixels. Each block is
transformed into DCT domain, which results in 8x8 coefficients.

The DCT process is used to remove the spatial redundancy between the pixels in
10 the same block. However, it does not consider the redundancy between the pixels from
different blocks. The first versions of the standards did not use any technique to exploit
the correlation between different blocks. Recently, MPEG-4 and H.263+ have added
tools/options to exploit this redundancy to certain extent. At present, MPEG-4 predicts
the DC coefficient (first coefficient, which is actually the block average) of the current
15 block by using the DC coefficients of the neighbouring blocks. H.263+ does this, and in
addition, it also predicts the first row or column of the DCT coefficients in some cases if
there is any benefit.

In brief, existing compression algorithms exploit the fact that the DCT coefficients
in the neighbouring blocks are sometimes similar to those in the current block. This
20 means that if the blocks contain completely different coefficients, the prediction will not
work.

Against this background, there is provided apparatus for coding video data,
comprising means for receiving pixel values organised in frames each comprising a
matrix of video blocks, each video block comprising a video matrix of N pixel values, and
25 processor means arranged to perform the following steps:

- a) to set each element in a prediction matrix to an initial prediction value;

- b) in the prediction matrix, to apply a smoothing transform to the values along the rows and then along the columns, or vice versa, to obtain interpolated values;
- c) to reset the prediction value to the interpolated value;
- d) to calculate the difference between the reset prediction values and
- 5 corresponding received pixel values to produce a residual prediction matrix containing the prediction residuals; and
- e) to perform a discrete cosine transform on the prediction residuals to obtain elements of a compressed video data matrix.

The processor means is preferably arranged iteratively to calculate the reset

10 prediction value used to calculate the prediction residual by repeating steps b) and c).

The number of iterations may be predetermined or, in an alternative, the iterations may be repeated until the change in the prediction value between one iteration and the next, is less than a predetermined threshold.

Step a) is most preferably performed by performing a discrete cosine transform on

15 the video matrix to obtain a transform video matrix of N coefficients, selecting n of the coefficients, setting the $N-n$ remaining coefficients to zero to obtain an initial prediction transform matrix of initial prediction coefficients, and performing an inverse discrete cosine transform on the initial prediction transform matrix to obtain a matrix of N initial prediction values.

20 In that case, the processor is preferably arranged to set n of the elements in the compressed video data matrix equal to the n coefficients selected from the transform video matrix, and to select the remaining $N-n$ coefficients from the prediction residuals.

The processor is further preferably arranged to adjust the prediction residuals before selecting the remaining $N - n$ elements, by:

- 25 f) performing a discrete cosine transform on the reset prediction value matrix to obtain a prediction transform matrix,
- g) selecting n coefficients from the transform prediction matrix,

h) subtracting the selected n transform prediction matrix coefficients from the selected n transform video coefficients to obtain n residual coefficients;

i) setting n elements of an adjustment transform matrix to the values of the n residual coefficients and setting $N - n$ remaining elements to zero;

5 j) performing an inverse discrete cosine transform on the adjustment transform matrix to obtain an adjustment value matrix; and

k) subtracting the adjustment value matrix from the reset prediction value matrix.

The apparatus may include means for processing pixels in a current and a
10 previous frame to produce pixel values which are the prediction residual between the actual pixel and a motion compensated pixel.

The invention extends to apparatus for expanding video data compressed by apparatus as claimed in any preceding claim, comprising means for receiving the compressed video matrix, and processor means arranged to perform the following steps:

15 a) to perform an inverse discrete cosine transform on received compressed video data to obtain a prediction residual matrix;

b) to set each element in a prediction matrix to the initial prediction value;

c) in the prediction matrix, to apply a smoothing transform to the values along the rows and then along the columns, or vice versa, to obtain interpolated values;

20 d) to reset the prediction value to the interpolated value; and

e) to calculate the sum of the reset prediction values and the prediction residual in corresponding positions in the received coded block matrix to produce an expanded video data matrix.

Embodiments of the invention will now be described, by way of example, with
25 reference to the accompanying drawings in which:

Figures 1A and 1B, when assembled as shown in Figure 1, show a block diagram of a transmitter including apparatus for compressing video data embodying the invention; and

Figures 2A and 2B, when assembled as shown in Figure 2, show a block diagram of a receiver including apparatus for expanding the video data compressed by the apparatus of Figure 1.

A frame of quantised and digitised pixel values is divided into video matrices comprising blocks of N pixels where as an example $N = 8 \times 8$. With a switches 1a, 1b set to "intra" as illustrated, a video matrix 2 is discrete cosine transformed in step 4 to produce a video transform matrix 6 comprising a block of N discrete cosine transform (DCT) coefficients where in the example $N = 8 \times 8$. Of these a square of n coefficients are selected in step 8, essentially the DC coefficient and optionally other coefficients.

In step 10, the remaining $N - n$ (i.e. $8 \times 8 - n$) coefficients are set to zero to obtain an initial prediction transform matrix 12. The coefficients are inverse discrete cosine transformed in step 14 to obtain an initial prediction matrix 16.

In step 18 interpolation is performed between the initial prediction values of matrix 16 and the values in the neighbouring preceding blocks to reset the prediction matrix. Values in a row 20, spatially nearest to the video matrix 2, are used in the interpolation process. Linear interpolation is performed between the value in a row/column position in the initial prediction matrix and the value in a corresponding column in row 20 weighted according to the distance in rows from the row 20.

Similarly values in a column 22, spatially nearest the video matrix 2, are used in the interpolation process. Linear interpolation is also performed between the value in a row/column position in the initial prediction matrix and the value in a corresponding row in column 22, weighted by the distance in columns from the column 22.

$$V_{\text{interpolated}} = \{2V_{r,c} + V_{20,c}/r + V_{r,22}/c\}^{1/4}$$

Where,

$V_{\text{interpolated}}$ is the interpolated prediction value, $V_{r,c}$ is the value at row r column c of the initial prediction matrix 16, $V_{20,c}$ is the value in column c of row 20, r is the distance in rows of the position r,C from row 20, $V_{r,22}$ is the value at row r in column 22, and C is the distance in columns of the position r,C from the column 22.

The interpolation step 18 is performed iteratively until, in one example, the change in values in one step is less than a predetermined threshold. In another example, a predetermined fixed number of iterations is performed.

When the iterations are complete, the reset prediction values are discrete cosine transformed in step 24 to obtain 8×8 coefficients of a transform prediction matrix 26. In step 28 n coefficients are selected and, in step 30 subtracted from the n video transform coefficients previously selected in step 8 to produce n residual coefficients. In step 32 the remaining $8 \times 8 - n$ coefficients are set to zero to obtain 8×8 adjustment coefficients 34. These are inverse discrete cosine transformed to produce 8×8 adjustment values.

The values of the reset prediction matrix are adjusted by subtracting from them the adjustment values. The values in the video matrix are subtracted from the adjusted reset prediction values to obtain a prediction residual matrix 34 of 8×8 values. In step 36, the prediction residual values are discrete cosine transformed to produce a transform residual matrix having 8×8 coefficients. Of these n will be zero because of the adjustment made to the reset prediction matrix.

The remaining $8 \times 8 - n$ coefficients are selected in step 38 and assembled with the n video transform coefficients previously selected in step 8 to provide a compressed video matrix of 8×8 coefficients. These are channel coded in step 40 and transmitted through a medium 42.

In the apparatus shown in Figures 2A and 2B, the signal received from the medium 42 is channel decoded in step 44 to produce a decoded compressed video data matrix 46 of 8×8 coefficients. Of these, n are selected in step 48 and the remaining $8 \times 8 - n$ are set to zero in step 50 to obtain a decoded initial prediction transform matrix 52
5 having 8×8 coefficients. The coefficients are inverse discrete cosine transformed to produce a +

decoded initial prediction matrix 54 having 8×8 initial prediction values.

In step 56, interpolation is performed iteratively on the initial prediction matrix in exactly the same manner as was performed in step 18 on the prediction matrix 16 using
10 the (decoded) neighbouring row 20 and column 22 to obtain a matrix 58 of reset prediction values.

In step 60, the remaining $8 \times 8 - n$ coefficients of matrix 46 are selected and n coefficients are set to zero in step 62 to obtain a decoded transform residual matrix 64 having 8×8 coefficients. These coefficients are inverse discrete cosine transformed in
15 step 66 to obtain a decoded prediction residual matrix having 8×8 residual values. In step 68 these are added to the reset prediction values in matrix 58 to produce a decoded video matrix 70 containing 8×8 pixel values corresponding to those of matrix 2.

Putting the switches 1a, 1b in their "inter" position, rearranges the apparatus to operate not on the current frame video matrix, but on the residual produced by
20 subtracting the values in a motion compensated block of a previous frame, from the values in the current frame video matrix 2 in step 74. The motion compensated values are added back in step 76 to produce the initial prediction matrix 16 values, and subtracted in step 78 from the reset prediction values.

In the expander shown in Figure 2, motion compensated values obtained in a decoded motion compensated video matrix 80 from a previously decoded frame, are added back in step 82 to produce the initial prediction matrix.

THIS PAGE BLANK (USPTO)

CLAIMS

1. Apparatus for coding video data, comprising means for receiving pixel values organised in frames each comprising a matrix of video blocks, each video block comprising a video matrix of N pixel values, and processor means arranged to perform
5 the following steps:

- a) to set each element in a prediction matrix to an initial prediction value;
- b) in the prediction matrix, to apply a smoothing transform to the values along the rows and then along the columns, or vice versa, to obtain interpolated values;
- c) to reset the prediction value to the interpolated value;
- 10 d) to calculate the difference between the reset prediction values and corresponding received pixel values to produce a residual prediction matrix containing the prediction residuals; and
- e) to perform a discrete cosine transform on the prediction residuals to obtain elements of a compressed video data matrix.

15 2. Apparatus as claimed in claim 1, wherein the processor means is arranged iteratively to calculate the reset prediction value used to calculate the prediction residual by repeating steps b) and c).

3. Apparatus as claimed in claim 2, wherein the number of iterations is predetermined.

20 4. Apparatus as claimed in claim 2, wherein the processor means is arranged to repeat the iterations until the change in the prediction value between one iteration and the next, is less than a predetermined threshold.

5. Apparatus as claimed in any preceding claim, wherein step a) is performed by performing a discrete cosine transform on the video matrix to obtain a
25 transform video matrix of N coefficients, selecting n of the coefficients, setting the $N-n$ remaining coefficients to zero to obtain an initial prediction transform matrix of initial

prediction coefficients, and performing an inverse discrete cosine transform on the initial prediction transform matrix to obtain a matrix of N initial prediction values.

6. Apparatus as claimed in claim 5, wherein the processor is arranged to set n of the elements in the compressed video data matrix equal to the n coefficients selected from the transform video matrix, and to select the remaining $N-n$ coefficients from the prediction residuals.

7. Apparatus as claimed in claim 6, wherein the processor is arranged to adjust the prediction residuals before selecting the remaining $N - n$ elements, by:

- f) performing a discrete cosine transform on the reset prediction value matrix to obtain a prediction transform matrix,
- g) selecting n coefficients from the transform prediction matrix,
- h) subtracting the selected n transform prediction matrix coefficients from the selected n transform video coefficients to obtain n residual coefficients;
- i) setting n elements of an adjustment transform matrix to the values of the n residual coefficients and setting $N - n$ remaining elements to zero;
- j) performing an inverse discrete cosine transform on the adjustment transform matrix to obtain an adjustment value matrix; and
- k) subtracting the adjustment value matrix from the reset prediction value matrix.

8. Apparatus as claimed in any preceding claim, including means for processing pixels in a current and a previous frame to produce pixel values which are the prediction residual between the actual pixel and a motion compensated pixel.

9. Apparatus for expanding video data compressed by apparatus as claimed in any preceding claim, comprising means for receiving the compressed video matrix, and processor means arranged to perform the following steps:

- a) to perform an inverse discrete cosine transform on received compressed video data to obtain a prediction residual matrix;

- b) to set each element in a prediction block matrix to the initial prediction value;
- c) in the prediction matrix, to apply a smoothing transform to the values along the rows and then along the columns, or vice versa, to obtain interpolated values;
- 5 d) to reset the prediction value to the interpolated value; and
- e) to calculate the sum of the reset prediction values and the prediction residual in corresponding positions in the received coded block matrix to produce an expanded video data matrix.

10. Apparatus as claimed in claim 8, wherein the processor means is
10 arranged iteratively to calculate the reset prediction value used to calculate the prediction residual by repeating steps b) and c).

11. Apparatus as claimed in claim 10, wherein the number of iterations is predetermined.

12. Apparatus as claimed in claim 10, wherein the processor means is
15 arranged to repeat the iterations until the change in the prediction value between one iteration and the next, is less than a predetermined threshold.

13. Apparatus as claimed in any of claims 9 to 12, wherein step a) is performed by performing a discrete cosine transform on the video matrix to obtain a transform video matrix of N coefficients, selecting n of the coefficients, setting the $N-n$
20 remaining coefficients to zero to obtain an initial prediction transform matrix of initial prediction coefficients, and performing an inverse discrete cosine transform on the initial prediction transform matrix to obtain a matrix of N initial prediction values.

14. Apparatus as claimed in claim 13 for expanding video data compressed by apparatus as claimed in claim 7, wherein the processor is arranged to select $N - n$
25 elements from the compressed video data matrix and to set n elements to zero before performing the inverse discrete cosine transform to obtain the prediction residual matrix.

THIS PAGE BLANK (USPTO)

ABSTRACT

APPARATUS FOR COMPRESSING AND EXPANDING VIDEO
DATA

Existing video data compression algorithms exploit the fact that the DCT
5 coefficients in the neighbouring blocks are sometimes similar to those in the current
block. This means that if the blocks contain completely different coefficients, the
prediction will not work.

Apparatus for coding video data is disclosed in which element in a prediction
matrix is set to an initial prediction value. In the prediction matrix, a smoothing transform
10 is applied to the values along the rows and then along the columns, or vice versa, to
obtain interpolated values. The prediction value is reset to the to the interpolated value
and the difference between the reset prediction values and corresponding received pixel
values is calculated to produce a residual prediction matrix containing the prediction
residuals. A discrete cosine transform is performed on the prediction residuals to obtain
15 elements of a compressed video data matrix.

The processor means is preferably arranged iteratively to calculate the reset
prediction value used to calculate the prediction residual by repeating steps b) and c).

20

25

Figure 1a and 1b

THIS PAGE BLANK (USPTO)

1/5

FIG. 1

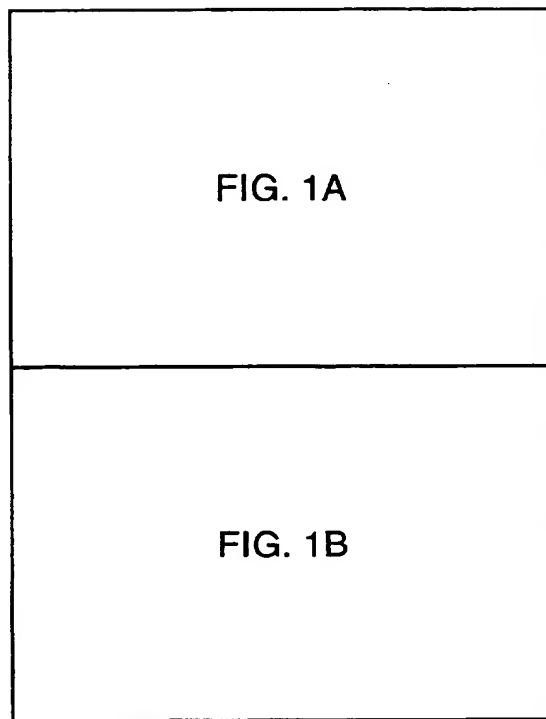
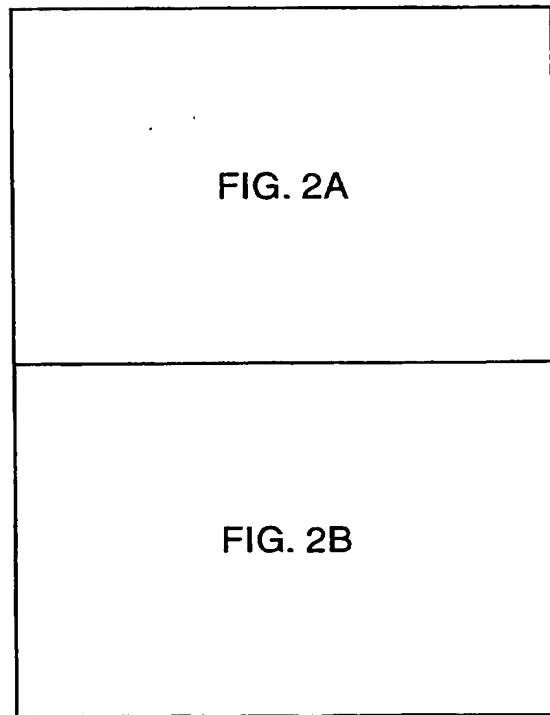


FIG. 2



CURRENT FRAME
VIDEO MATRIX
8 x 8 PIXELS



3/5

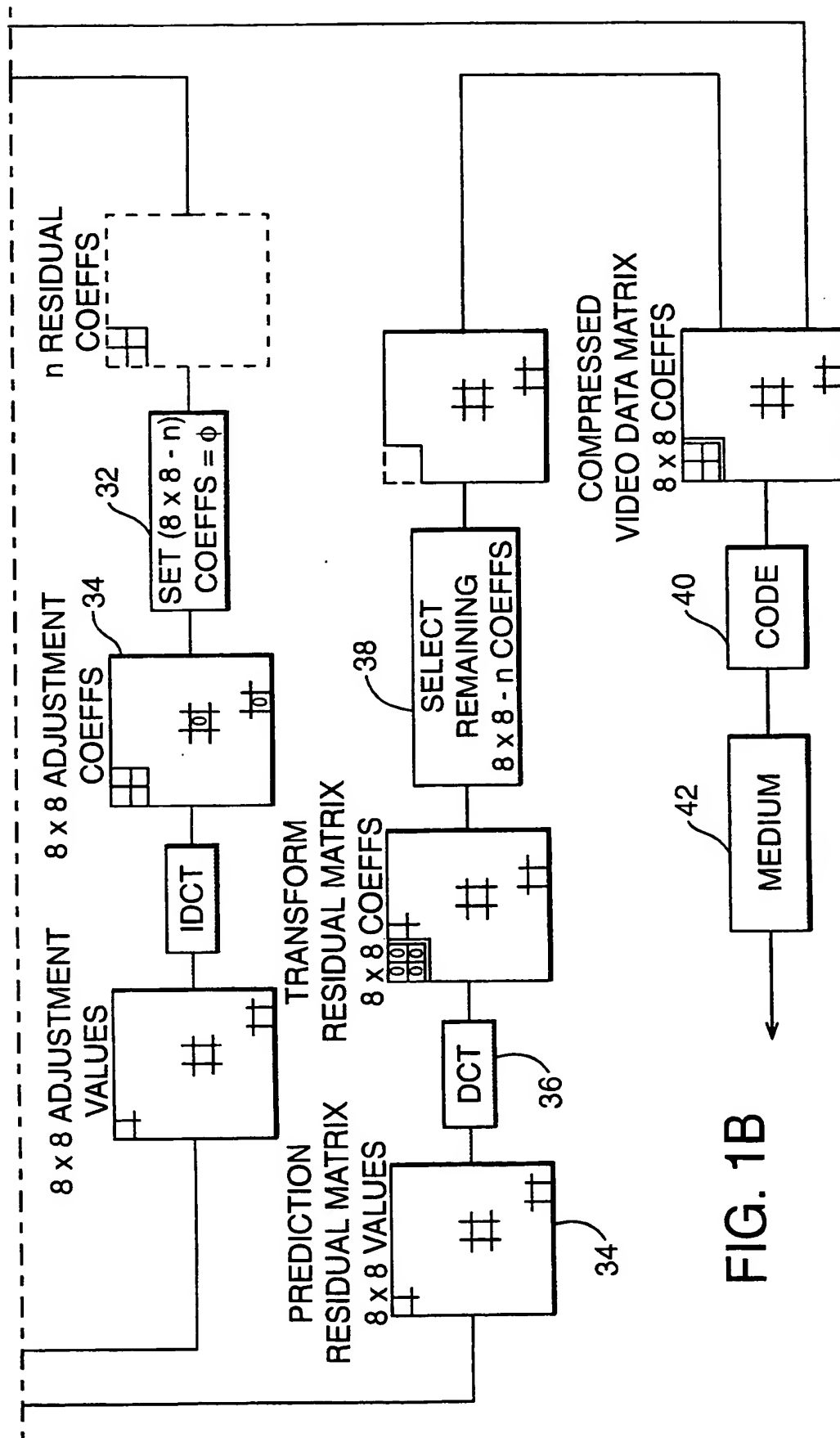
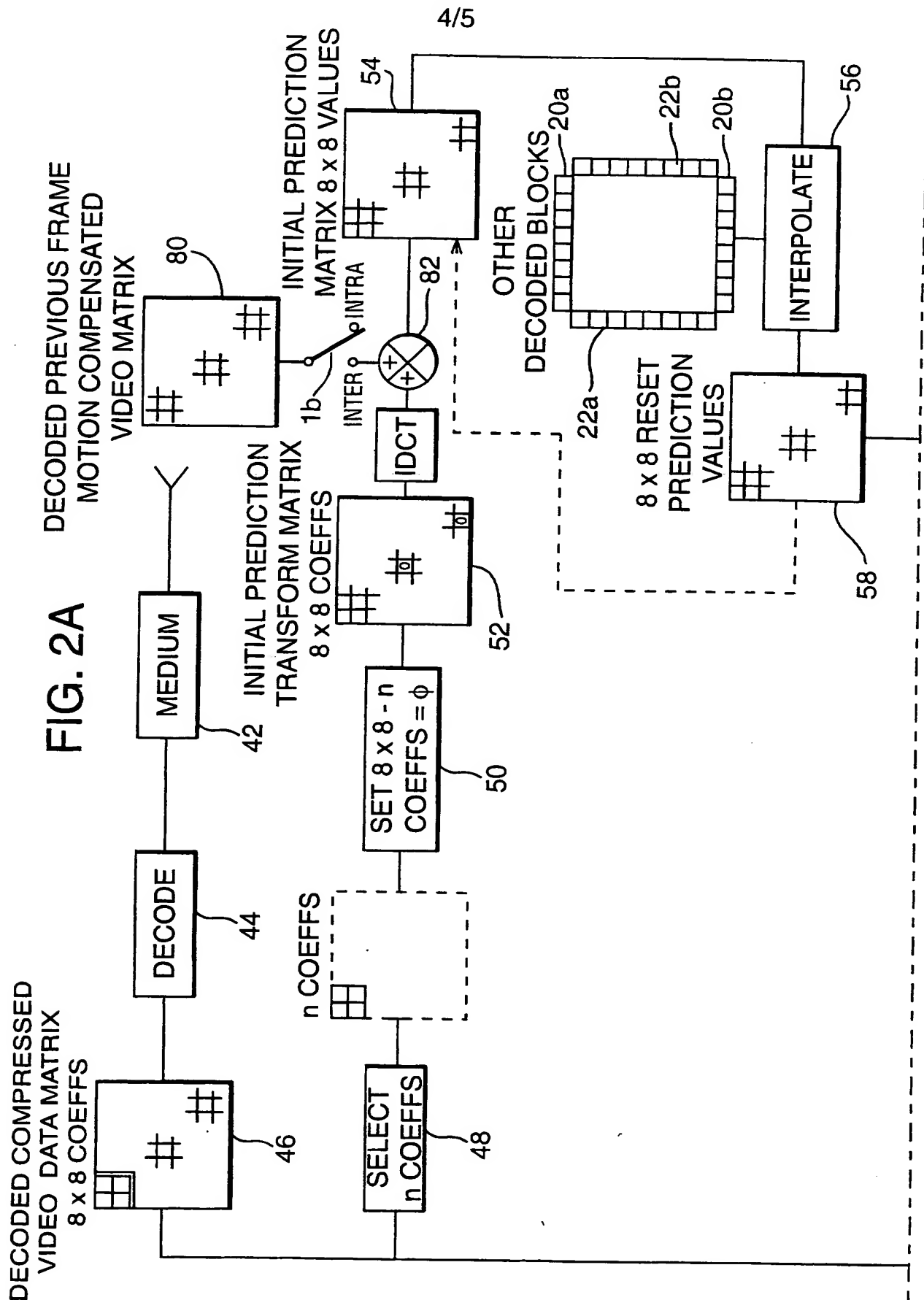


FIG. 1B

DECODED COMPRESSED
VIDEO DATA MATRIX
8 x 8 COEFFS



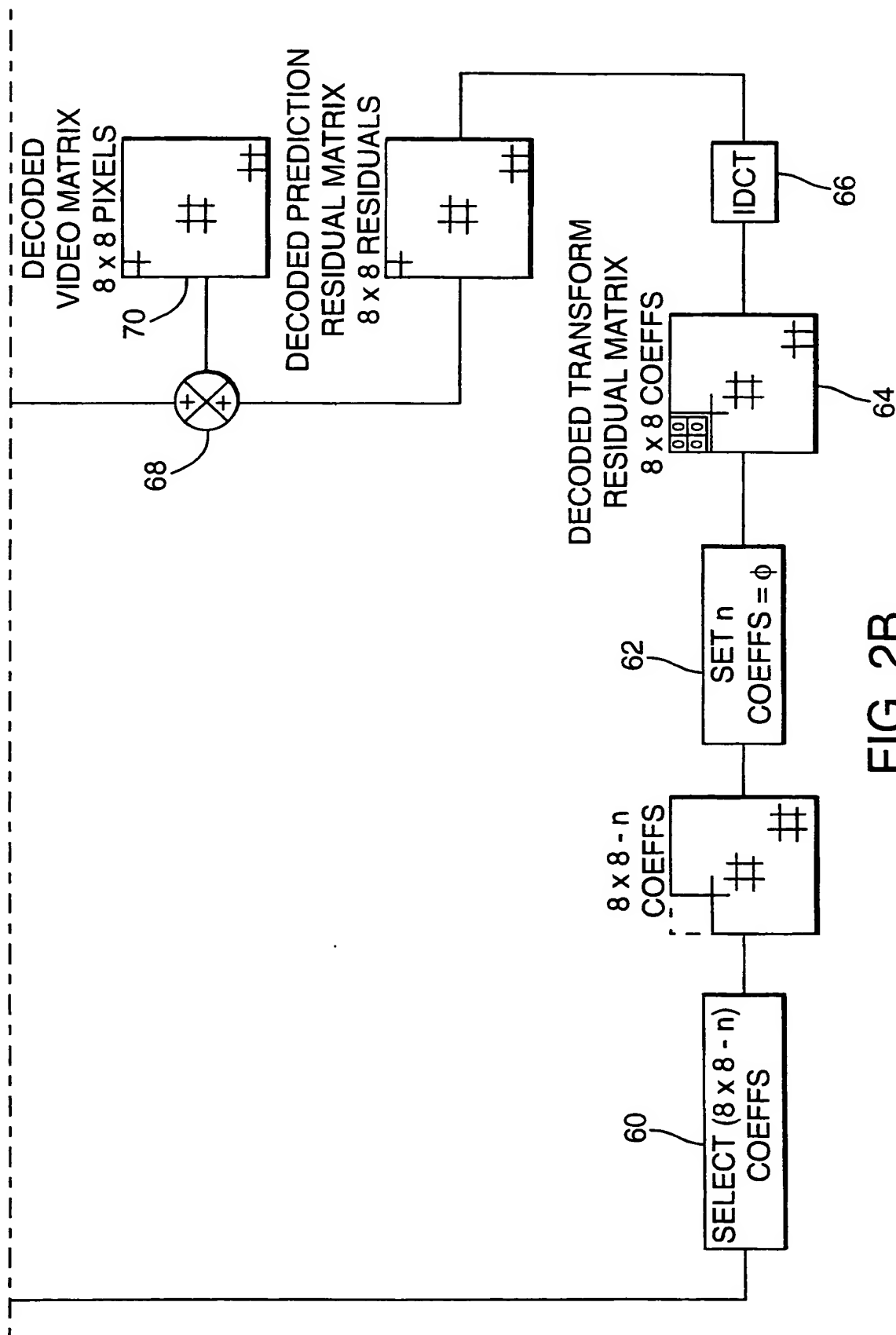


FIG. 2B

THIS PAGE BLANK (USPTO)